DMI, 26-27 October 2016

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

EXCELENCIA
SEVERO
OCHOA

EC-Earth3.2 at the BSC-ES

François Massonnet

and the Climate Prediction Group

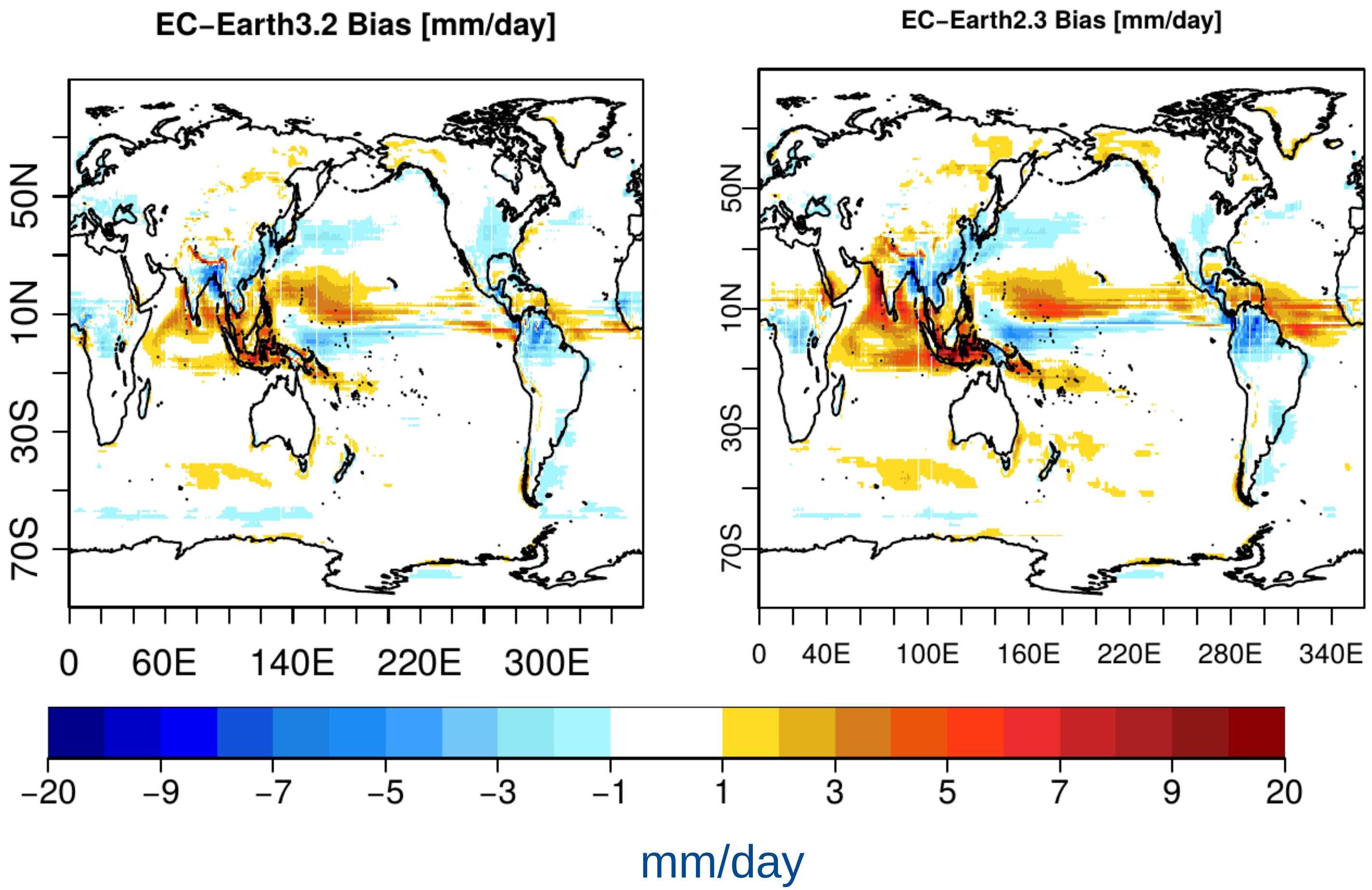
R. Bilbao, O. Bellprat L. Brodeau, R. Cruz-García, F. Doblas-Reyes, E. Exarchou, N. Fučkar, J. García-Serrano, V. Guemas, M. Ménégoz, C. Prodhomme, V. Sicardi, E. Tourigny

Status of EC-Earth3.2 at BSC

- ORCA1L75
  - Runs in seasonal prediction mode
  - Stratospheric aerosols are being added (M. Ménégoz)
- ORCA025L75
  - Many compatibility issues were fixed (see portal): bathymetry, closed seas, …
  - Does not run beyond 20 time steps, trying with short time step.
- PRIMAVERA & HighResMIP formatting: ongoing (E. Tourigny)
- Reproducibility tests are underway (Massonnet, Ménégoz, Acosta)
- More to follow at the EC-Earth meeting

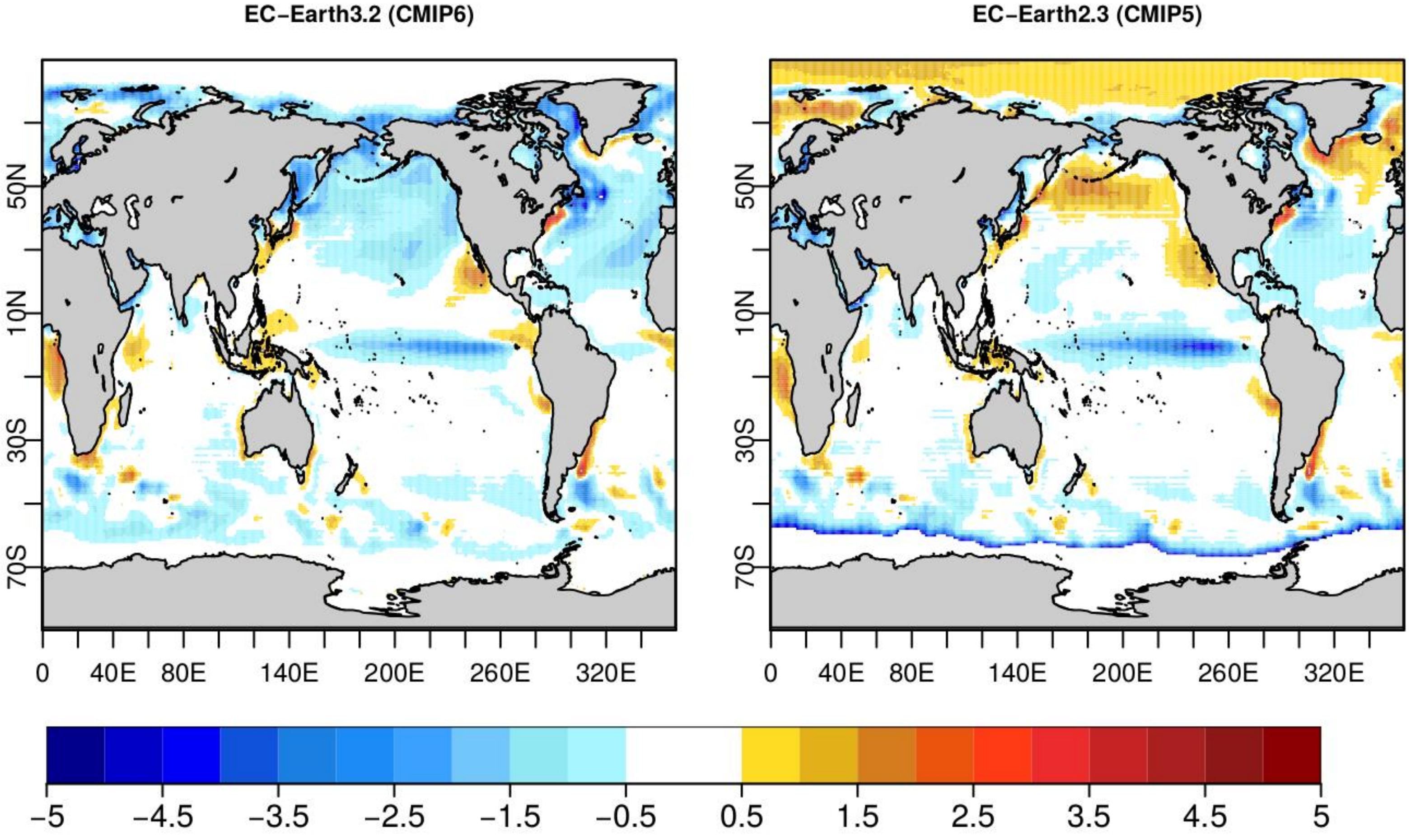
JJA Precipitation bias (1993-2009)

CMIP6 CMIP5


EC−Earth3.2 Bias [mm/day]
EC−Earth2.3 Bias [mm/day]
50N
10N
30S
70S
0
60E
140E
220E
300E
0
40E
100E
160E
220E
280E
340E
−20
−9
−7
−5
−3
−1
1
3
5
7
9
20
mm/day

JJA SST 1993-2009 bias

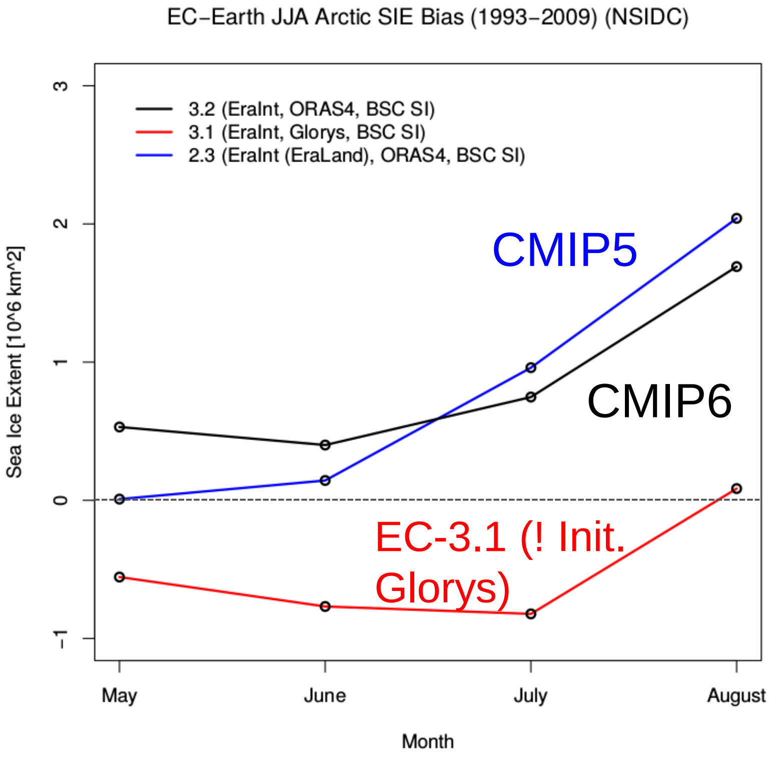
CMIP6 CMIP5


EC−Earth3.2 (CMIP6)
EC−Earth2.3 (CMIP5)
50N
10N
30S
70S
0
40E
80E
140E
200E
260E
320E
−5
−4.5
−3.5
−2.5
−1.5
−0.5
0.5
1.5
2.5
3.5
4.5
5

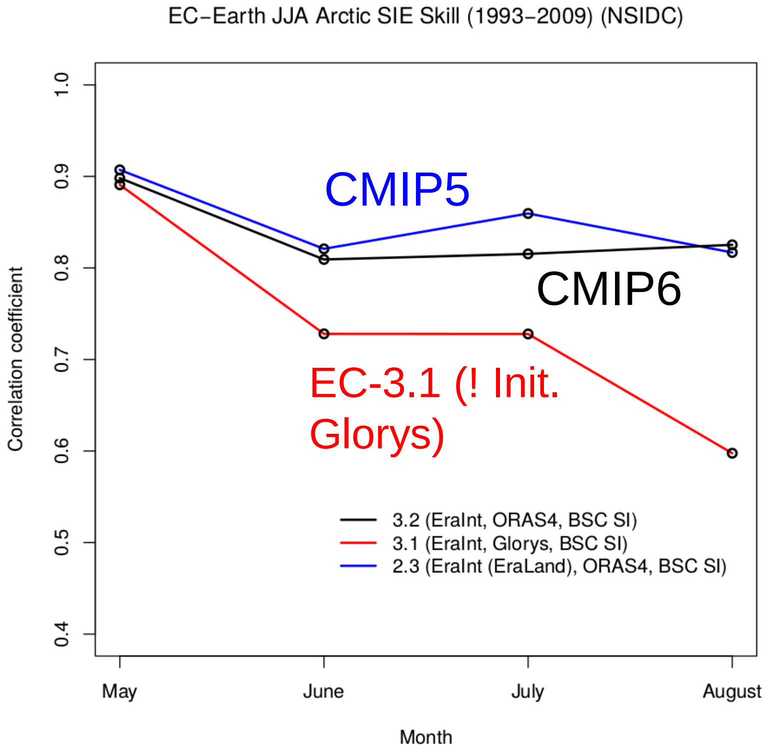
°C

JJA Arctic sea ice extent 1993-2009

Bias

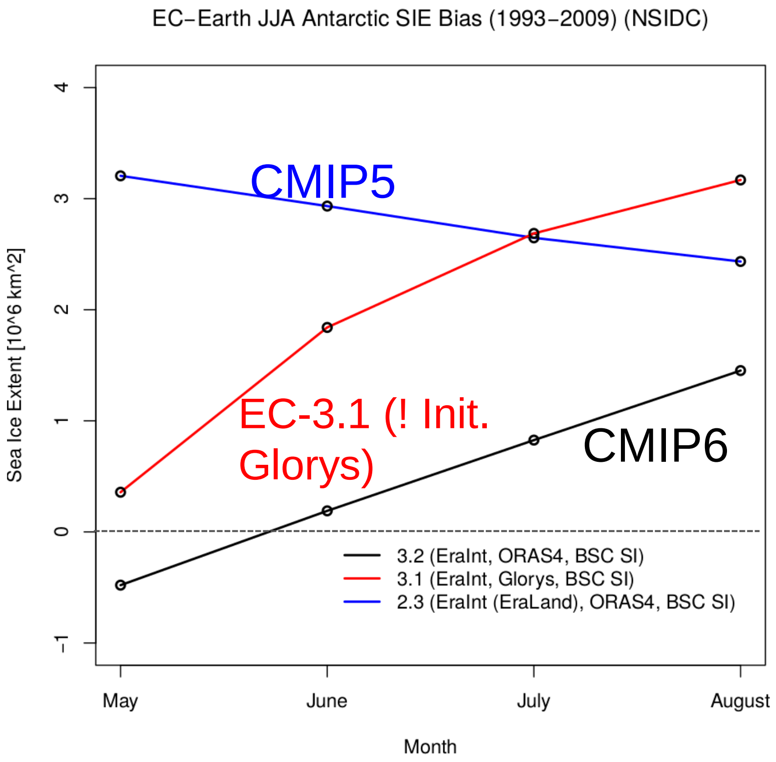


Correlation

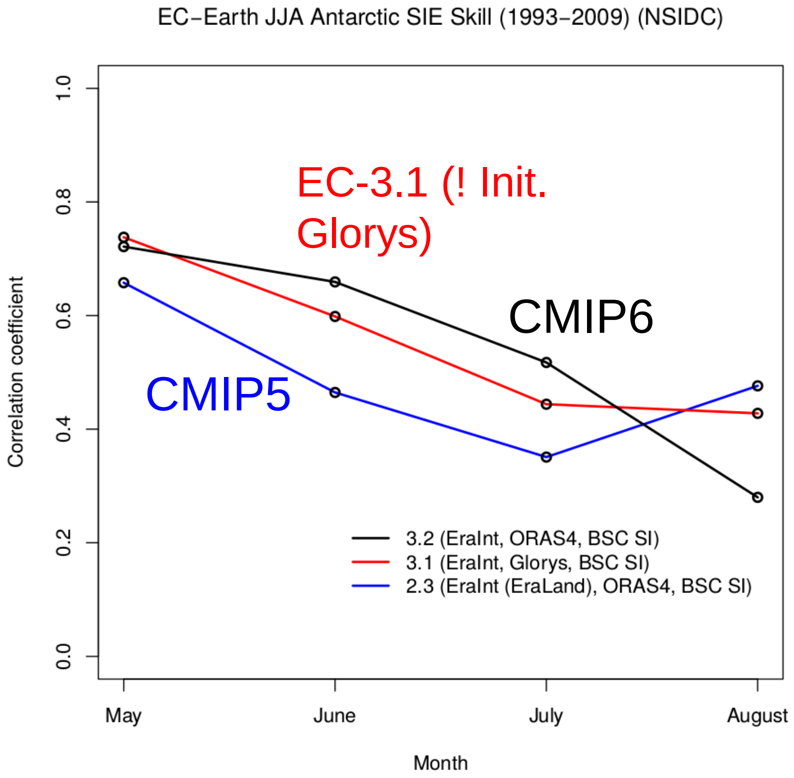


JJA Antarctic sea ice extent 1993-2009

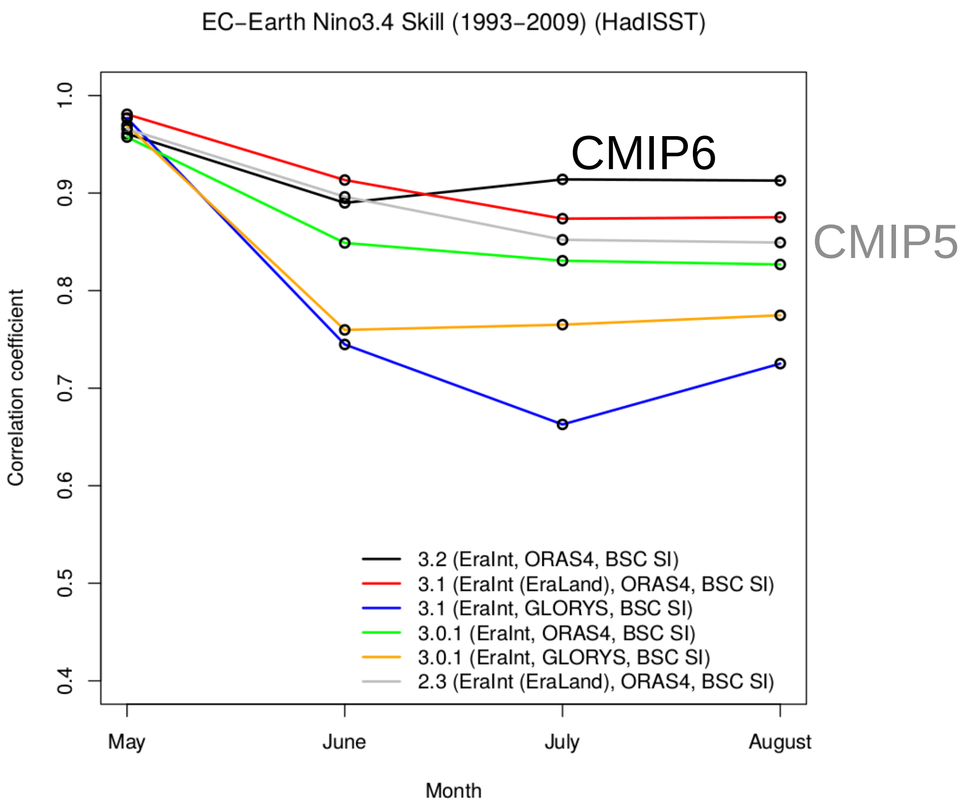
Bias



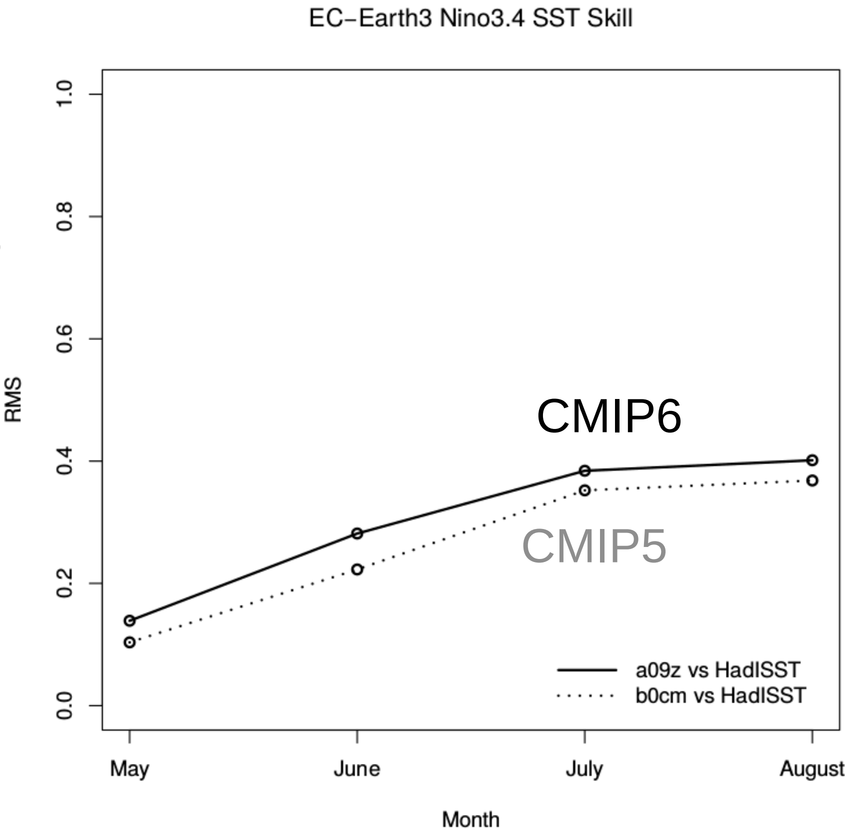
Correlation



JJA Niño3.4 SST 1993-2009



EC−Earth Nino3.4 Skill (1993−2009) (HadISST)



EC−Earth3 Nino3.4 SST Skill

DMI, 26-27 October 2016

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

EXCELENCIA
SEVERO
OCHOA

The EnKF in EC-Earth3.2

François Massonnet

and the Climate Prediction Group

R. Bilbao, O. Bellprat L. Brodeau, R. Cruz-García, F. Doblas-Reyes, E. Exarchou, N. Fučkar, J. García-Serrano, V. Guemas, M. Ménégoz, C. Prodhomme, V. Sicardi, E. Tourigny

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

What is an EnKF?

1. What is the scientific purpose of an EnKF?

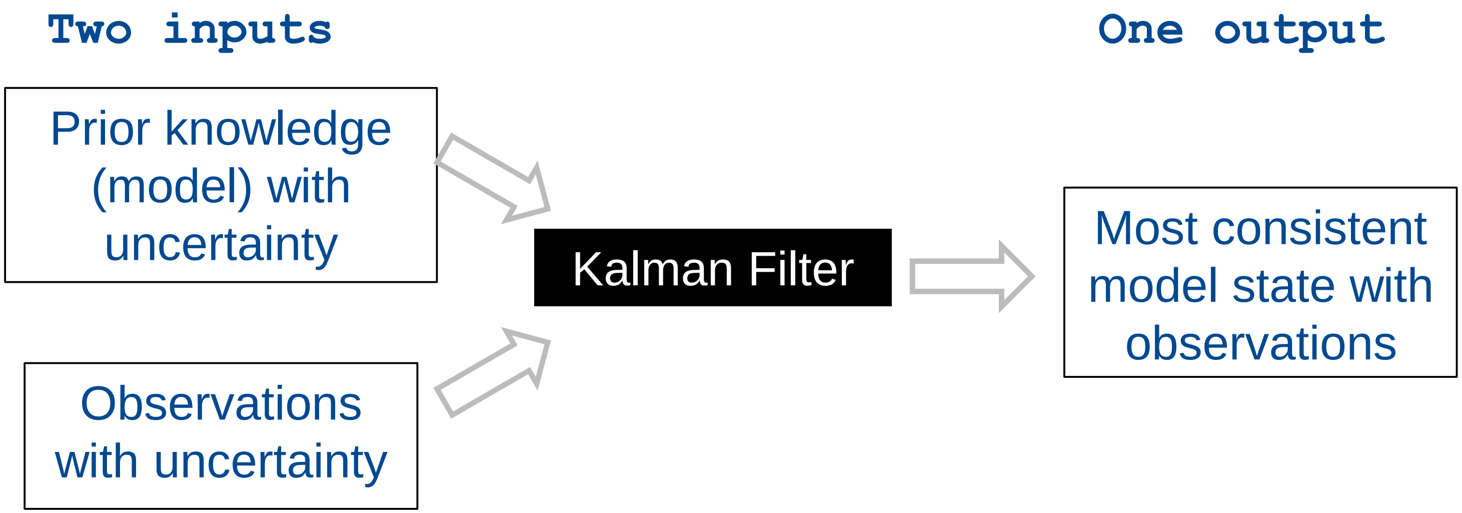
A Kalman Filter (KF) is a multivariate statistical problem of state estimation by maximizing the posterior state distribution, given observations, assuming bias-free models and observations and gaussian distribution of errors.

Every word is important!!!

What is an EnKF?

1. What is the scientific purpose of an EnKF?

A Kalman Filter (KF) is a multivariate statistical problem of state estimation by maximizing the posterior state distribution, given observations, assuming bias-free models and observations and gaussian distribution of errors.



Two inputs

One output

Prior knowledge (model) with uncertainty

Observations with uncertainty

Kalman Filter

Most consistent model state with observations

An **EnKF** works as a KF, but uses **ensembles** to estimate prior uncertainty

What is an EnKF?

2. What are possible applications of the EnKF in climate science?

A. State estimation for reconstruction (reanalyses)
B. State estimation for initialization (initial conditions)
C. Parameter estimation (tuning/calibration)

Three « hot » questions of climate science are addressed with the same tool! How cool is that?

This was the subject of my thesis, see references.

What is an EnKF?

3. What are the advantages and downsides w.r.t. other methods?

	Nudging	Variational (3D/4D Var)	Sequential methods (EnKF)	Particle filtering
Ease of implementation	Fair	Hard (coding adjoint)	Medium-Hard	Medium
CPU consumption	Low	Low-Medium	High (~20-50 members)	Very high (~100 members to avoid degenerate solution)
Needs changes in model code?	Yes (restoring term added to tendencies)	Yes (adjoint)	No	No
Physical consistency	Little (univariate)	Up to linear approximation	Up to linear approximation	Full consistency
Estimation of prior uncertainty	None	Static	Dynamic	Dynamic
Assumptions to reach optimality	Not defined as an optimization problem	Gaussian centered errors	Gaussian centered errors	None
Produces ensembles?	No	No	Yes (hence available as IC)	Yes (hence available as IC)

The EnKF is, all in all, a honest tradeoff to estimate a system state with satisfactory physical consistency given limited resources available

Frequently asked questions

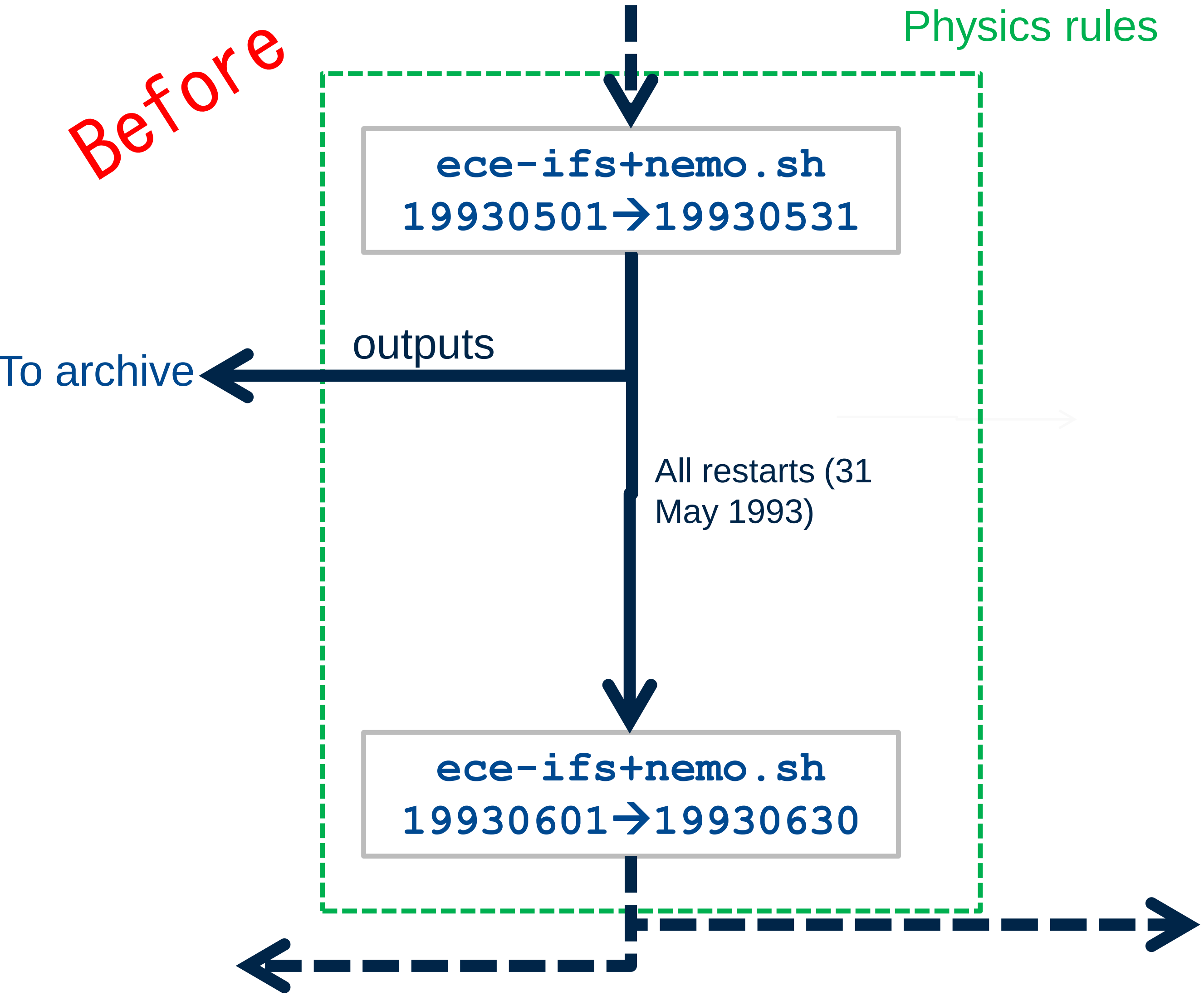
1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

How does an EnKF work?

1. How does it affect the workflow?



Before

Physics rules

`ece-ifs+nemo.sh`
`19930501→19930531`

outputs

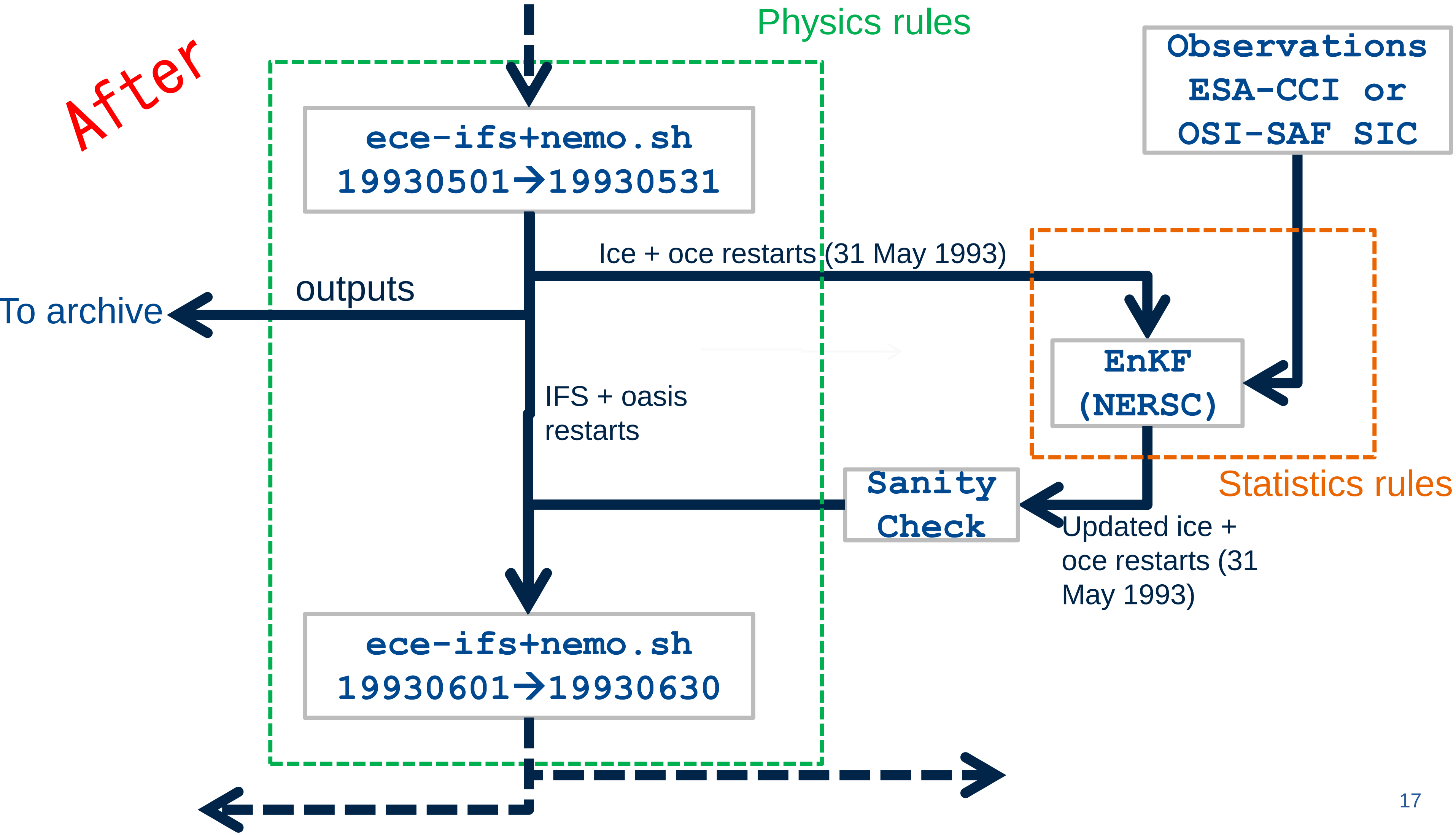
To archive

All restarts (31 May 1993)

`ece-ifs+nemo.sh`
`19930601→19930630`

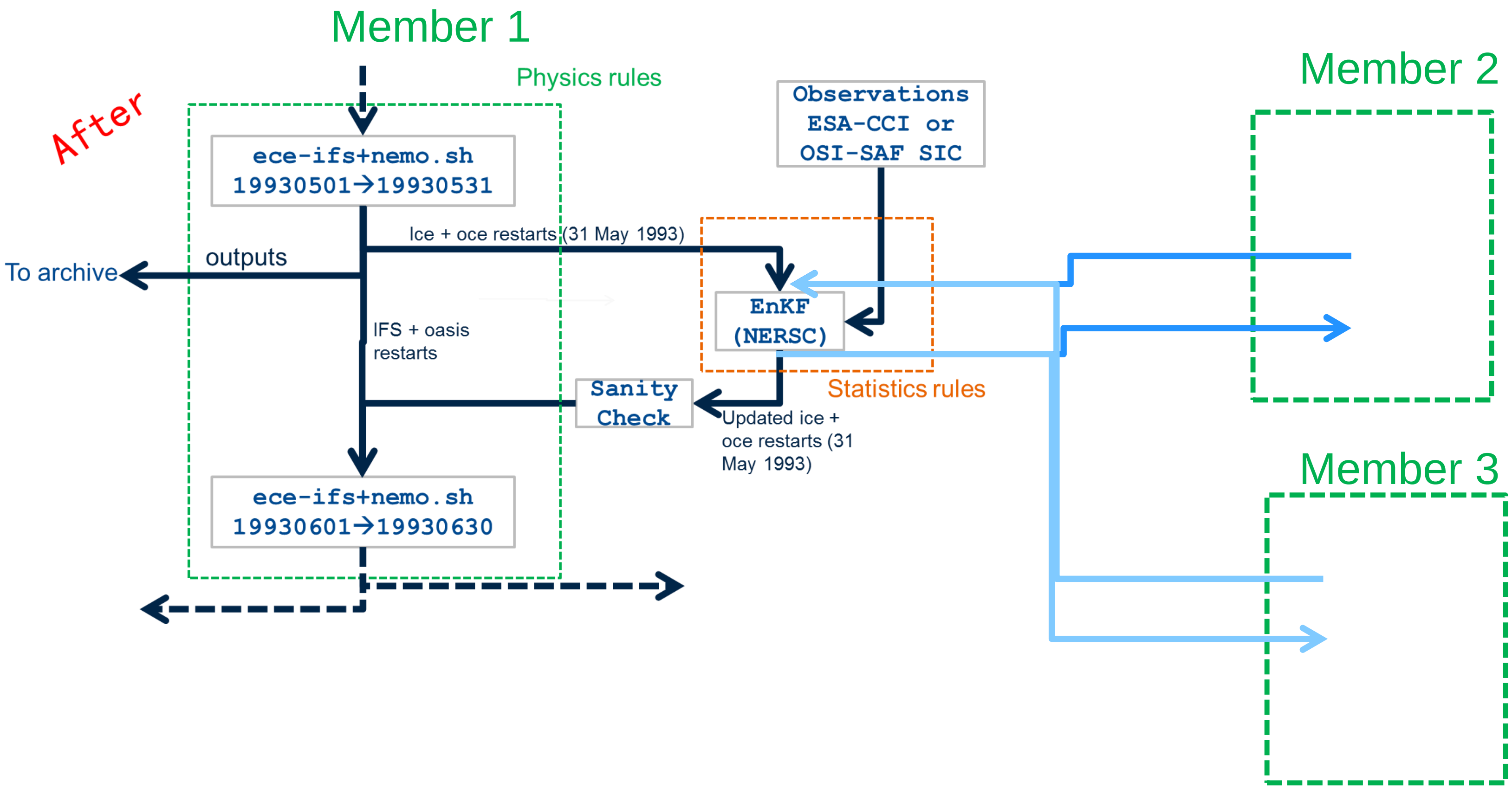
How does an EnKF work?

1. How does it affect the workflow?

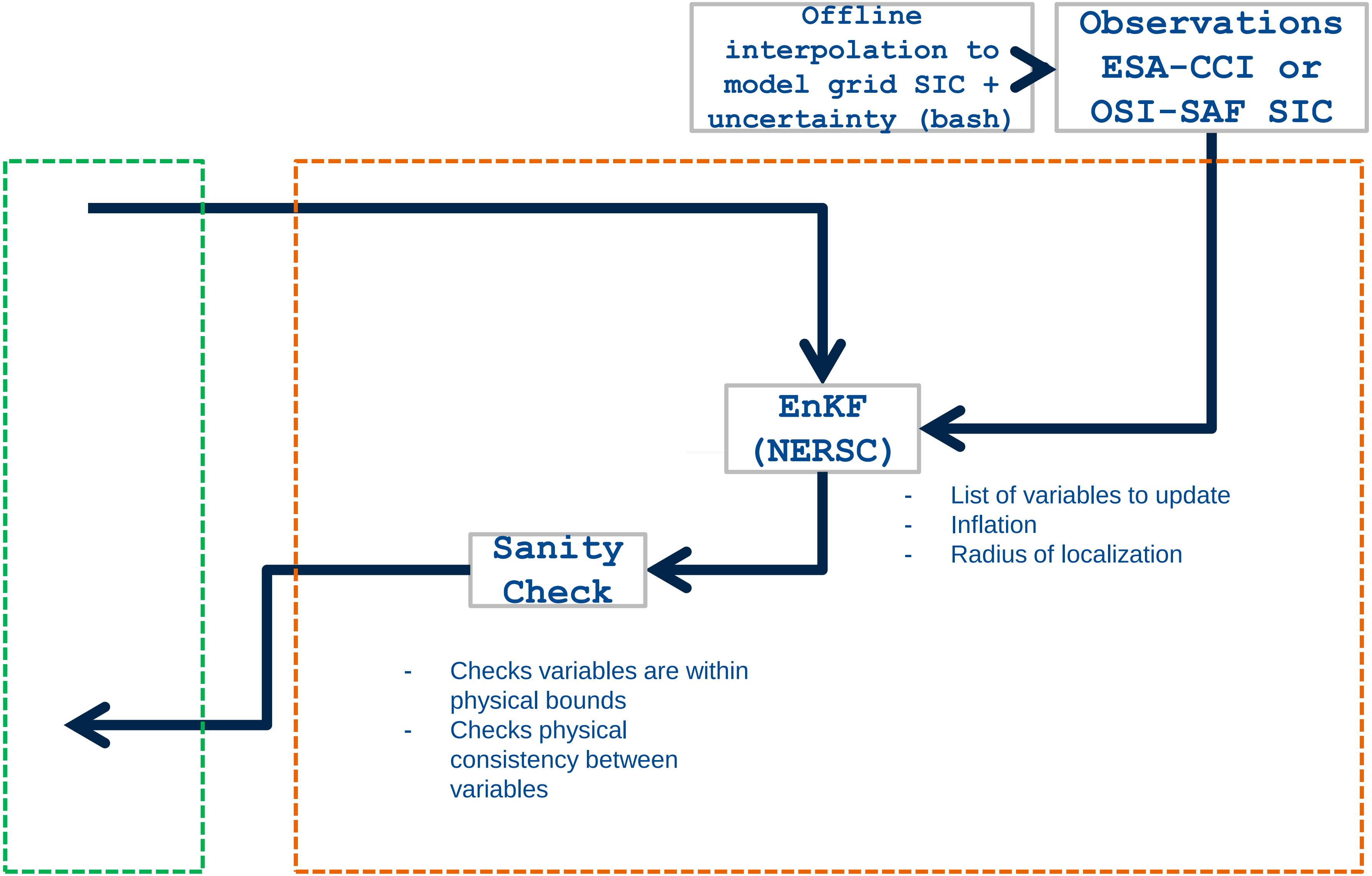


After
Physics rules
ece-ifs+nemo.sh
19930501→19930531
Observations
ESA-CCI or
OSI-SAF SIC
Ice + oce restarts (31 May 1993)
outputs
To archive
EnKF
(NERSC)
IFS + oasis
restarts
Sanity
Check
Statistics rules
Updated ice +
oce restarts (31
May 1993)
ece-ifs+nemo.sh
19930601→19930630

How does an EnKF work?


Member 1
Physics rules
After
ece-ifs+nemo.sh
19930501→19930531
To archive
outputs
Ice + oce restarts (31 May 1993)
IFS + oasis
restarts
Observations
ESA-CCI or
OSI-SAF SIC
EnKF
(NERSC)
Statistics rules
Sanity
Check
Updated ice +
oce restarts (31
May 1993)
ece-ifs+nemo.sh
19930601→19930630
Member 2
Member 3

How does an EnKF work?



Offline interpolation to model grid SIC + uncertainty (bash)

Observations ESA-CCI or OSI-SAF SIC

EnKF (NERSC)

- List of variables to update
- Inflation
- Radius of localization

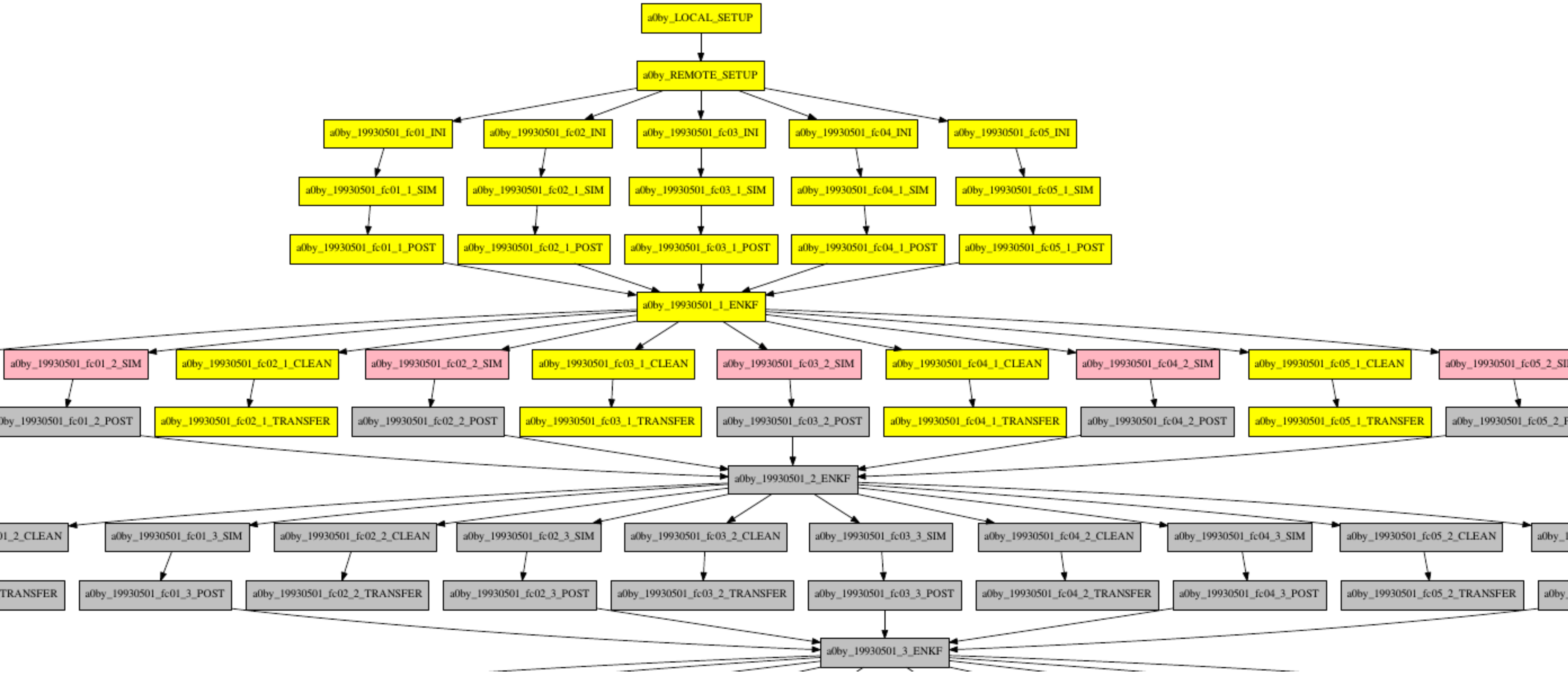
Sanity Check

- Checks variables are within physical bounds
- Checks physical consistency between variables

Statistics rules

How does an EnKF work?

Implementation in a workflow manager: Autosubmit


a0by_LOCAL_SETUP
a0by_REMOTE_SETUP
a0by_19930501_fc01_INI
a0by_19930501_fc02_INI
a0by_19930501_fc03_INI
a0by_19930501_fc04_INI
a0by_19930501_fc05_INI
a0by_19930501_fc01_1_SIM
a0by_19930501_fc02_1_SIM
a0by_19930501_fc03_1_SIM
a0by_19930501_fc04_1_SIM
a0by_19930501_fc05_1_SIM
a0by_19930501_fc01_1_POST
a0by_19930501_fc02_1_POST
a0by_19930501_fc03_1_POST
a0by_19930501_fc04_1_POST
a0by_19930501_fc05_1_POST
a0by_19930501_1_ENKF
a0by_19930501_fc01_2_SIM
a0by_19930501_fc02_1_CLEAN
a0by_19930501_fc02_2_SIM
a0by_19930501_fc03_1_CLEAN
a0by_19930501_fc03_2_SIM
a0by_19930501_fc04_1_CLEAN
a0by_19930501_fc04_2_SIM
a0by_19930501_fc05_1_CLEAN
a0by_19930501_fc05_2_SI
0by_19930501_fc01_2_POST
a0by_19930501_fc02_1_TRANSFER
a0by_19930501_fc02_2_POST
a0by_19930501_fc03_1_TRANSFER
a0by_19930501_fc03_2_POST
a0by_19930501_fc04_1_TRANSFER
a0by_19930501_fc04_2_POST
a0by_19930501_fc05_1_TRANSFER
a0by_19930501_fc05_2_F
a0by_19930501_2_ENKF
01_2_CLEAN
a0by_19930501_fc01_3_SIM
a0by_19930501_fc02_2_CLEAN
a0by_19930501_fc02_3_SIM
a0by_19930501_fc03_2_CLEAN
a0by_19930501_fc03_3_SIM
a0by_19930501_fc04_2_CLEAN
a0by_19930501_fc04_3_SIM
a0by_19930501_fc05_2_CLEAN
a0by_
TRANSFER
a0by_19930501_fc01_3_POST
a0by_19930501_fc02_2_TRANSFER
a0by_19930501_fc02_3_POST
a0by_19930501_fc03_2_TRANSFER
a0by_19930501_fc03_3_POST
a0by_19930501_fc04_2_TRANSFER
a0by_19930501_fc04_3_POST
a0by_19930501_fc05_2_TRANSFER
a0by_
a0by_19930501_3_ENKF

How does an EnKF work?

2. Is it dependent on the version of EC-Earth used?

NO, unless state variables are added, deleted or renamed in EC-Earth

3. What happens to the ocean when sea ice is assimilated

To cut a long story short, all oceanic variables and sea ice variables are updated even though only SIC is assimilated.

The update of each variable (T, S, U, V, …) is based on the relationship it has with SIC, learned from the ensemble.

This guarantees a state that is consistent up to the linear approximation.

Frequently asked questions

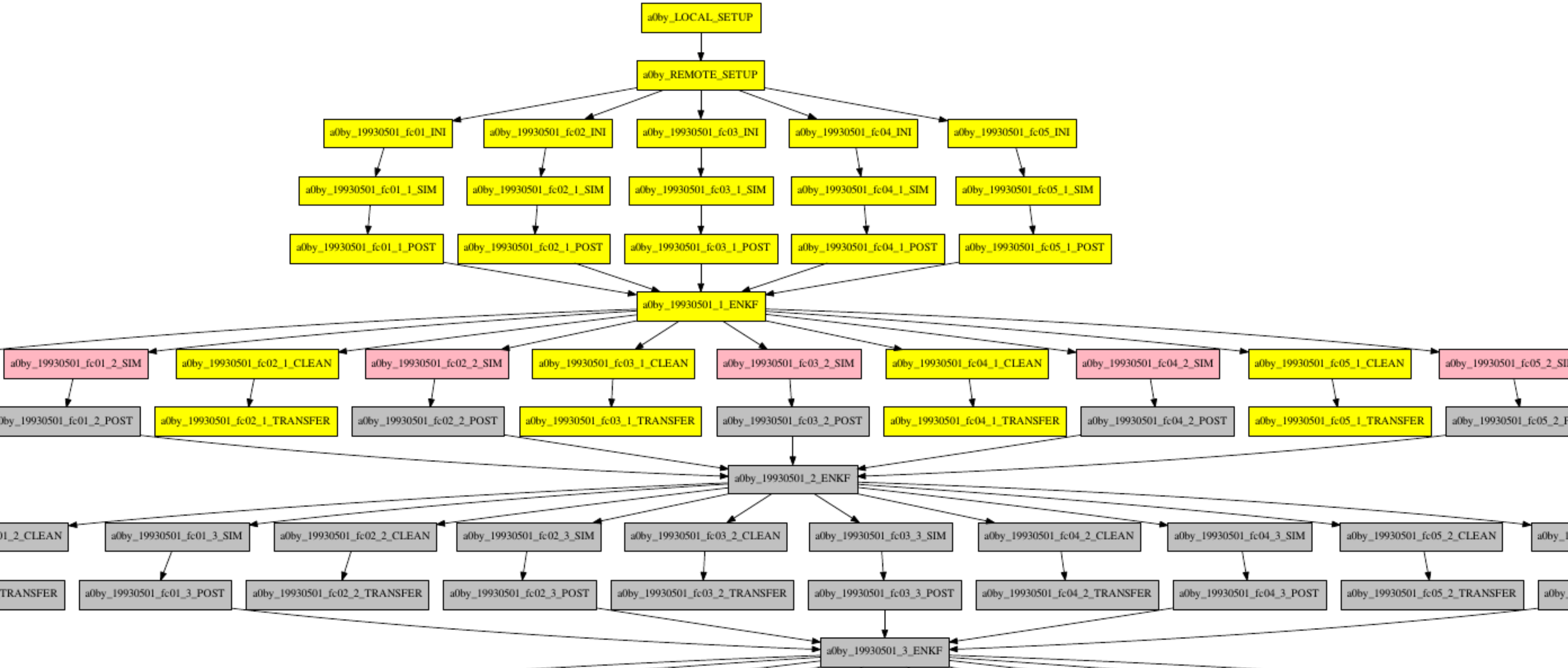
1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

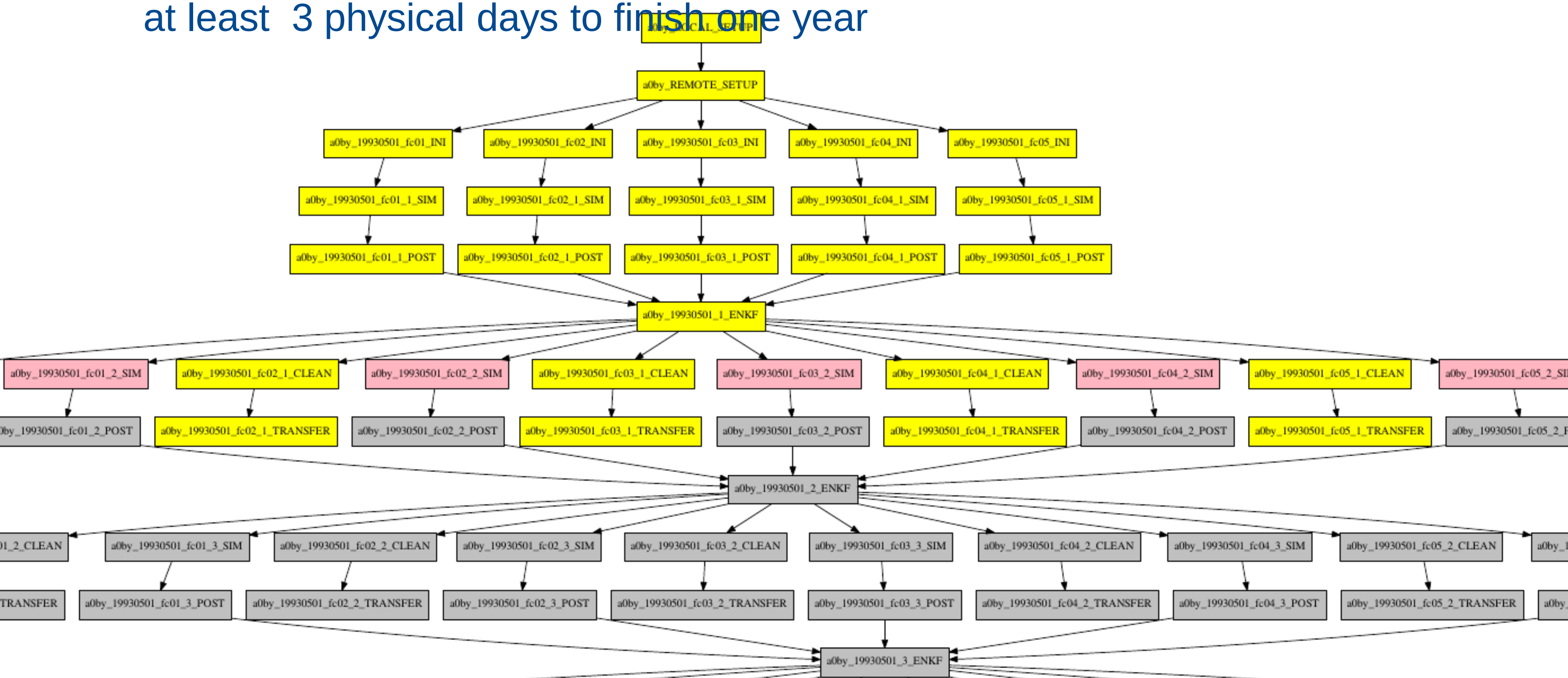
How much extra resources does the EnKF require?

- There are N=24 members to integrate, hence 24 SIM jobs, but all 24 jobs can be submitted independently.
- The EnKF job is itself quite fast (15 min on 16 procs) but cannot start until the last of the 24 SIM jobs is finished ☹. That’s the bottleneck of sequential methods.


a0by_LOCAL_SETUP
a0by_REMOTE_SETUP
a0by_19930501_fc01_INI
a0by_19930501_fc02_INI
a0by_19930501_fc03_INI
a0by_19930501_fc04_INI
a0by_19930501_fc05_INI
a0by_19930501_fc01_1_SIM
a0by_19930501_fc02_1_SIM
a0by_19930501_fc03_1_SIM
a0by_19930501_fc04_1_SIM
a0by_19930501_fc05_1_SIM
a0by_19930501_fc01_1_POST
a0by_19930501_fc02_1_POST
a0by_19930501_fc03_1_POST
a0by_19930501_fc04_1_POST
a0by_19930501_fc05_1_POST
a0by_19930501_1_ENKF
a0by_19930501_fc01_2_SIM
a0by_19930501_fc02_1_CLEAN
a0by_19930501_fc02_2_SIM
a0by_19930501_fc03_1_CLEAN
a0by_19930501_fc03_2_SIM
a0by_19930501_fc04_1_CLEAN
a0by_19930501_fc04_2_SIM
a0by_19930501_fc05_1_CLEAN
a0by_19930501_fc05_2_SI
0by_19930501_fc01_2_POST
a0by_19930501_fc02_1_TRANSFER
a0by_19930501_fc02_2_POST
a0by_19930501_fc03_1_TRANSFER
a0by_19930501_fc03_2_POST
a0by_19930501_fc04_1_TRANSFER
a0by_19930501_fc04_2_POST
a0by_19930501_fc05_1_TRANSFER
a0by_19930501_fc05_2_F
a0by_19930501_2_ENKF
01_2_CLEAN
a0by_19930501_fc01_3_SIM
a0by_19930501_fc02_2_CLEAN
a0by_19930501_fc02_3_SIM
a0by_19930501_fc03_2_CLEAN
a0by_19930501_fc03_3_SIM
a0by_19930501_fc04_2_CLEAN
a0by_19930501_fc04_3_SIM
a0by_19930501_fc05_2_CLEAN
a0by_1
TRANSFER
a0by_19930501_fc01_3_POST
a0by_19930501_fc02_2_TRANSFER
a0by_19930501_fc02_3_POST
a0by_19930501_fc03_2_TRANSFER
a0by_19930501_fc03_3_POST
a0by_19930501_fc04_2_TRANSFER
a0by_19930501_fc04_3_POST
a0by_19930501_fc05_2_TRANSFER
a0by_
a0by_19930501_3_ENKF

How much extra resources does the EnKF require?

- At ORCA1L75 with EC-Earth3.2:
  - CPU consumption is ~ 24 [members] x 432 [procs] x 0.2 [hrs] = **2 kCPUhrs per cycle.** (one cycle = one month here). This makes **~0.8 million CPUhrs for producing initial conditions over a full period.** CPU consumed during EnKF jobs is negligible.
  - At ORCA1, with EC-Earth3.2, it takes between 2 hours and 5 hours of physical time (depending on load on machine) to complete a full SIM-EnKF cycle. Expect at least 3 physical days to finish one year


a0by_LOCAL_SETUP
a0by_REMOTE_SETUP
a0by_19930501_fc01_INI
a0by_19930501_fc02_INI
a0by_19930501_fc03_INI
a0by_19930501_fc04_INI
a0by_19930501_fc05_INI
a0by_19930501_fc01_1_SIM
a0by_19930501_fc02_1_SIM
a0by_19930501_fc03_1_SIM
a0by_19930501_fc04_1_SIM
a0by_19930501_fc05_1_SIM
a0by_19930501_fc01_1_POST
a0by_19930501_fc02_1_POST
a0by_19930501_fc03_1_POST
a0by_19930501_fc04_1_POST
a0by_19930501_fc05_1_POST
a0by_19930501_1_ENKF
a0by_19930501_fc01_2_SIM
a0by_19930501_fc02_1_CLEAN
a0by_19930501_fc02_2_SIM
a0by_19930501_fc03_1_CLEAN
a0by_19930501_fc03_2_SIM
a0by_19930501_fc04_1_CLEAN
a0by_19930501_fc04_2_SIM
a0by_19930501_fc05_1_CLEAN
a0by_19930501_fc05_2_SI
0by_19930501_fc01_2_POST
a0by_19930501_fc02_1_TRANSFER
a0by_19930501_fc02_2_POST
a0by_19930501_fc03_1_TRANSFER
a0by_19930501_fc03_2_POST
a0by_19930501_fc04_1_TRANSFER
a0by_19930501_fc04_2_POST
a0by_19930501_fc05_1_TRANSFER
a0by_19930501_fc05_2_F
a0by_19930501_2_ENKF
01_2_CLEAN
a0by_19930501_fc01_3_SIM
a0by_19930501_fc02_2_CLEAN
a0by_19930501_fc02_3_SIM
a0by_19930501_fc03_2_CLEAN
a0by_19930501_fc03_3_SIM
a0by_19930501_fc04_2_CLEAN
a0by_19930501_fc04_3_SIM
a0by_19930501_fc05_2_CLEAN
a0by_1
TRANSFER
a0by_19930501_fc01_3_POST
a0by_19930501_fc02_2_TRANSFER
a0by_19930501_fc02_3_POST
a0by_19930501_fc03_2_TRANSFER
a0by_19930501_fc03_3_POST
a0by_19930501_fc04_2_TRANSFER
a0by_19930501_fc04_3_POST
a0by_19930501_fc05_2_TRANSFER
a0by_
a0by_19930501_3_ENKF

How much extra resources does the EnKF require?

Disk space:

- If you want to save everything

- 24 members, save original (forecast) and updated (analysis) sea ice and ocean restarts + IFS + OASIS restarts: 5.5 Go per member, 132 Go per cycle.
- This makes 55 To for a full period (35 yr, monthly cycles)

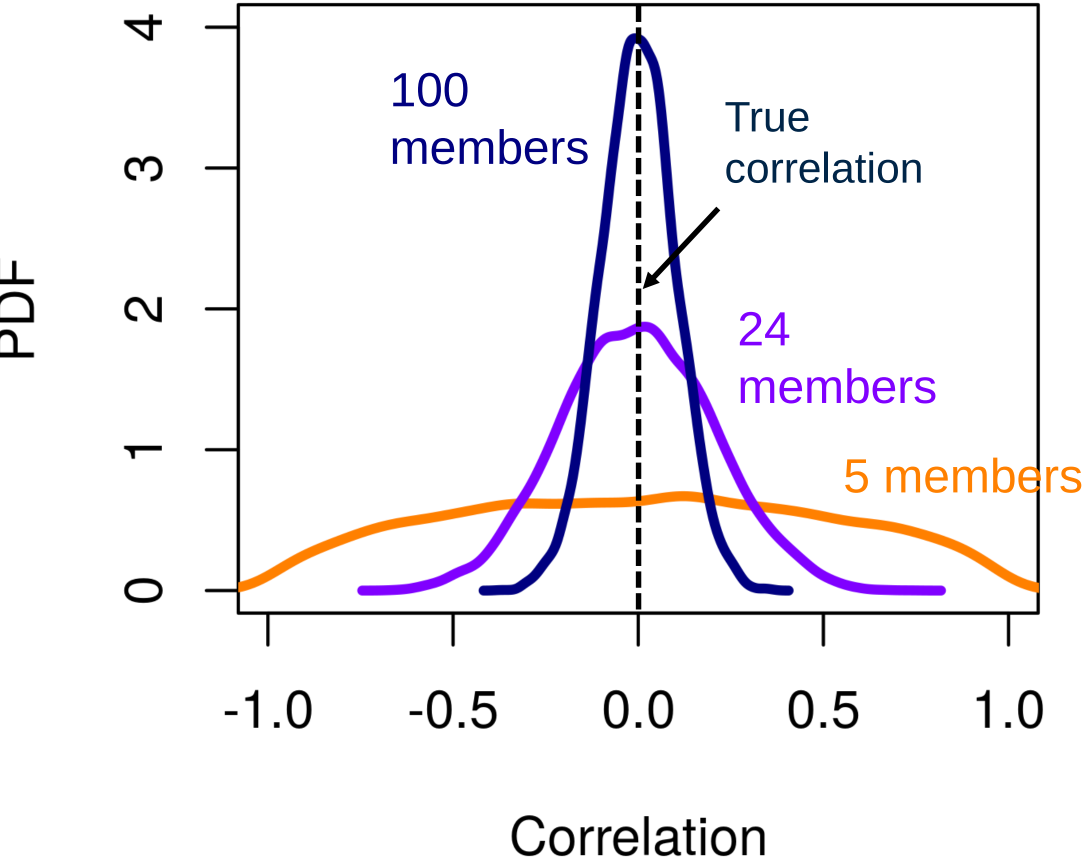
- If you want to save only the analyzed restarts (not forecasts), 40 To.

- If you want to save only the mean of all 24 restarts, this makes 1.7 To. But then, an alternative perturbation strategy has to be proposed when starting predictions from the initial states produced by the EnKF

How much extra resources does the EnKF require?

- Do we really need 24 members? Can't we just run 5?

PDF of sample correlation between 2 variables with true correlation = 0.0

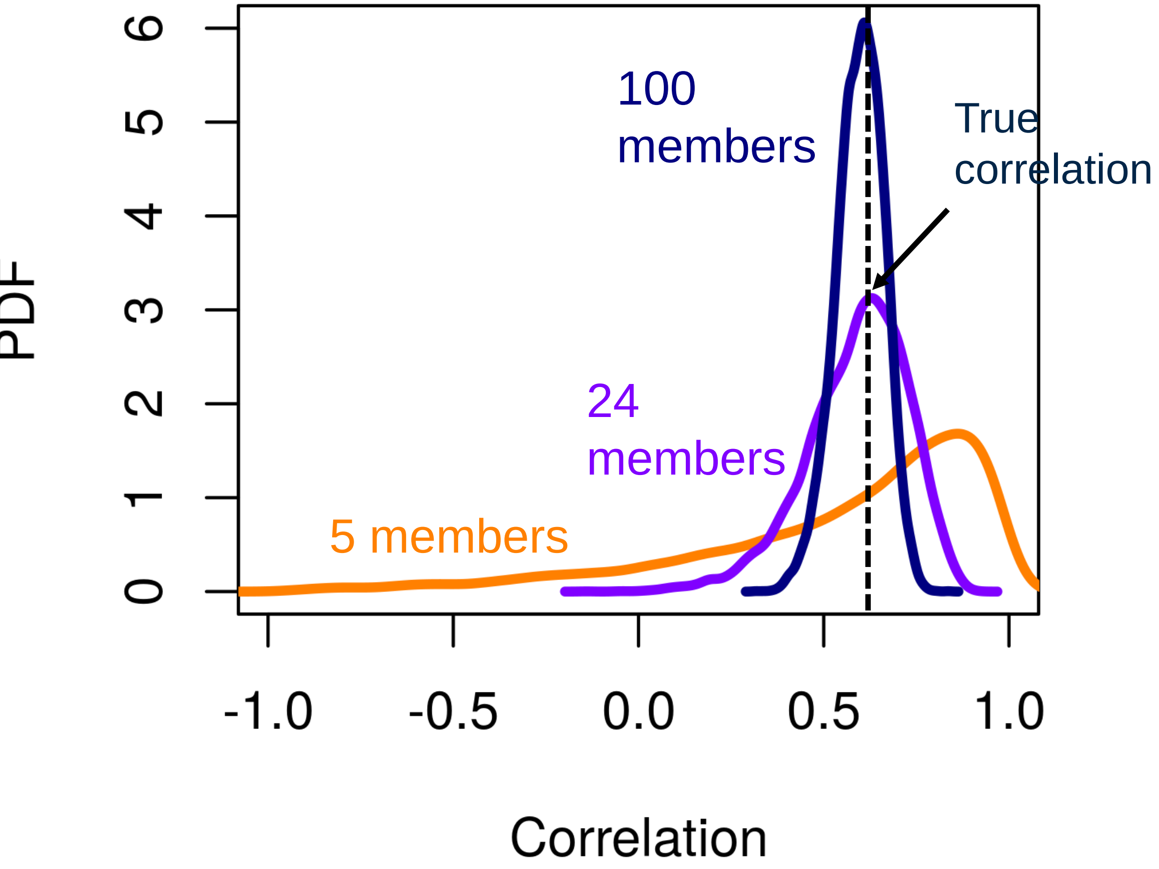


With 5 members, there are good chances to observe spurious correlations even though no physical link is actually present

How much extra resources does the EnKF require?

- Do we really need 24 members? Can't we run just 5?

PDF of sample correlation between 2 variables with true correlation = 0.6



If a physical link exists (e.g. SSS-SIC), then with 5 members we will wrongly estimate the strength of the association, leading to either too strong or too weak updates.

How much extra resources does the EnKF require?

- Can we call the EnKF routine less often than every month?

That’s a test to do.

But the longest the assimilation window, the less constrained the initial state.

Also, the longest the assimilation window, the larger the spread, the larger the bias, and hence the larger the updates, the larger the risk to have restartability issues (because all updates are done in a linear space)

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

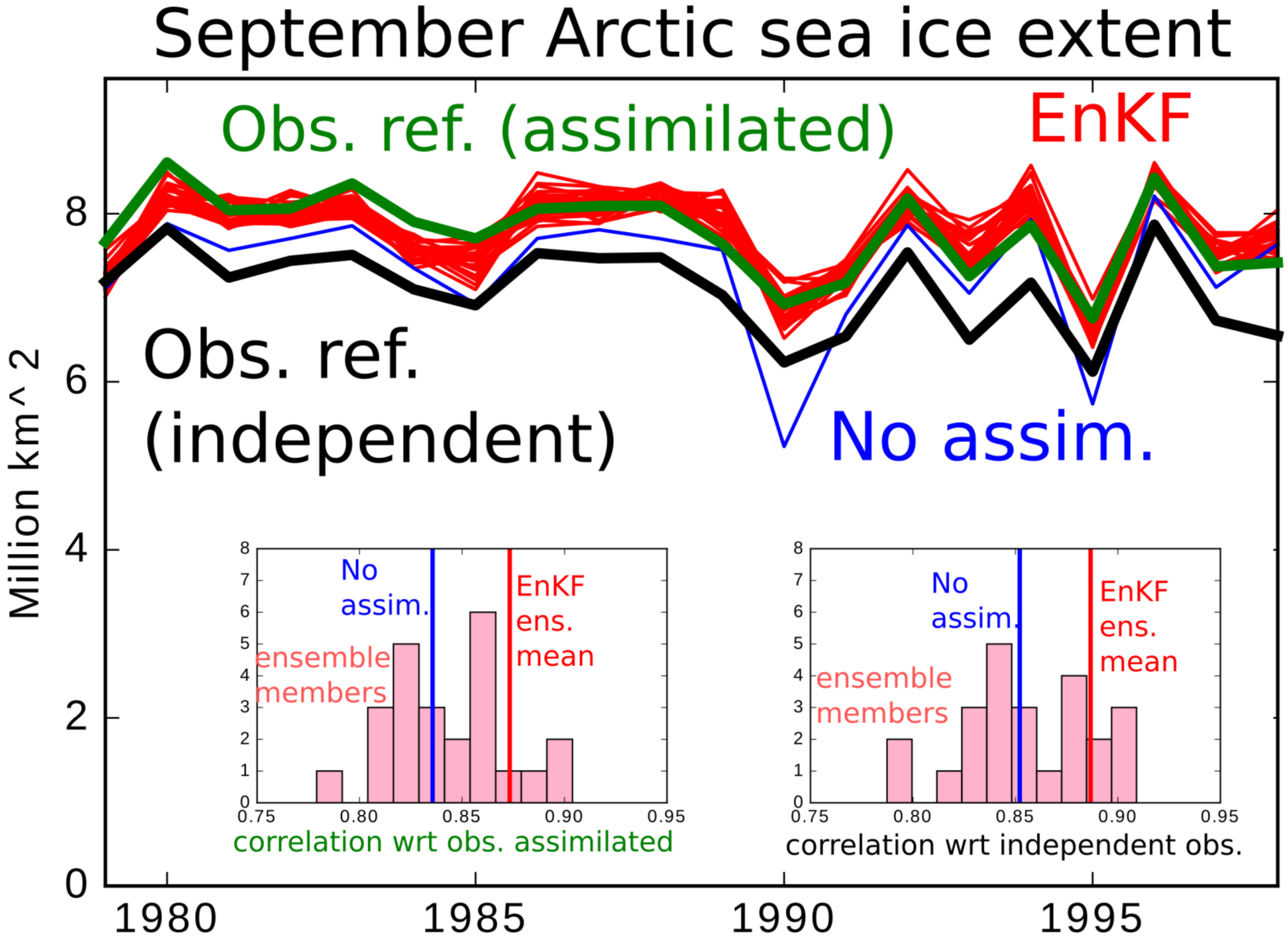
What is the status of the EnKF in EC-Earth?

Two configurations are currently working at BSC-ES

- Stand-alone simulations
  - NEMO3.6, forced by perturbed Drakkar Forcing Set 5.2
  - 24 members
  - 1979-1999 completed
  - Assimilation of OSI-SAF SIC (1979-1993) and ESA-CCI SIC (1993-1999).
  - Simulation is ongoing on MareNostrum3

- Coupled simulations
  - EC-Earth3.2
  - 24 members
  - 1993-1994 completed
  - Assimilation of ESA-CCI SIC (1993-1999)
  - Simulation is ongoing on MareNostrum3

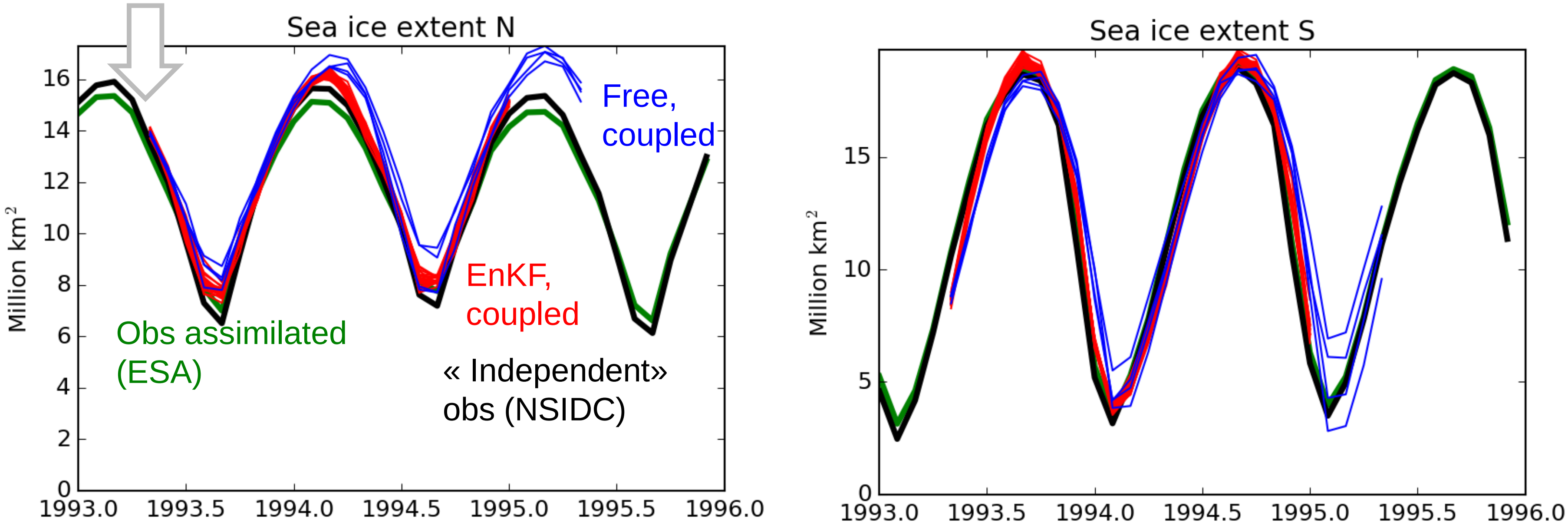
Results from stand-alone config.



Results from coupled experiments

The EnKF is able to keep EC-Earth track the observed sea ice extent despite systematic biases in the model

Start from IC obtained in EnKF forced reanalysis



Free, coupled

EnKF, coupled

Obs assimilated (ESA)

« Independent» obs (NSIDC)

The model is over confident in winter in Arctic, we may try to assimilate less often to increase the magnitude of the updates

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
    1. What is the scientific purpose of an EnKF?
    2. What are possible applications of the EnKF in climate sciences?
    3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
    1. How does it affect the workflow?
    2. Is it dependent on the version of EC-Earth used?
    3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
    1. What is the typical length of an assimilation cycle?
    2. What is the space required to save the experiment?
    3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

Frequently asked questions

1. What is an ensemble Kalman Filter (EnKF)?
   1. What is the scientific purpose of an EnKF?
   2. What are possible applications of the EnKF in climate sciences?
   3. What are the advantages and downsides wrt other methods?
2. How does an EnKF work?
   1. How does it affect the workflow?
   2. Is it dependent on the version of EC-Earth used?
   3. What happens with the ocean when sea ice is assimilated?
3. How much extra resources does the EnKF require?
   1. What is the typical length of an assimilation cycle?
   2. What is the space required to save the experiment?
   3. Can we run with less than 25 members?
4. What is the status of the EnKF in EC-Earth?
5. What are the plans of the EnKF in EC-Earth?

What are the plans of the EnKF in EC-Earth?

- (short-term) The ongoing stand-alone and coupled reanalyses will be extended to cover a full 1979-now period

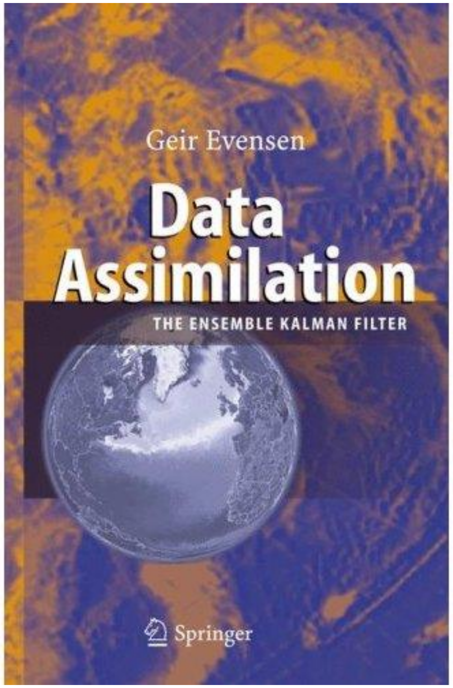
- (short-term) The ongoing stand-alone reanalysis will be the base for DCPP
- (short-term) An assessment of the quality of the reanalysis will be done

- (mid-term): launch **ORCA025 stand-alone** EnKF

- (long-term): change EnKF code to also update atmosphere

- (long-term): launch **ORCA025 coupled** EnKF

References



Geir Evensen

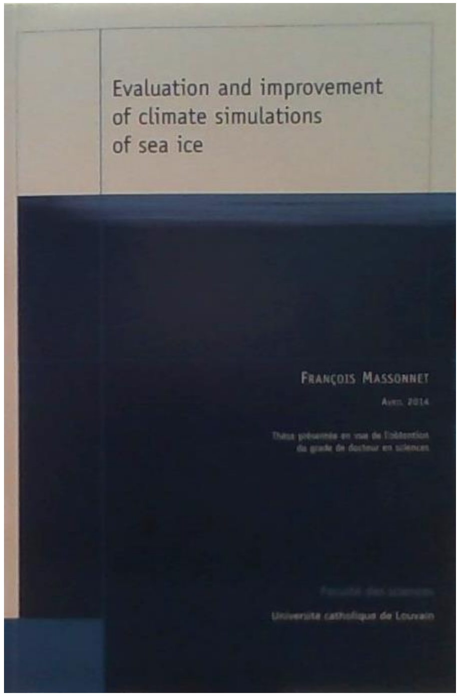
Data Assimilation

THE ENSEMBLE KALMAN FILTER

Springer

Evensen’s book
(basics, theory,
examples)

http://enkf.nersc.no/
(source code)



Evaluation and improvement
of climate simulations
of sea ice

François Massonnet

Avril 2014

Université catholique de Louvain

F. Massonnet’s thesis
(applications to sea ice)

www.climate.be/u/fmasson